EXHIBIT R

TO DECLARATION OF S. MERRILL WEISS IN SUPPORT OF PLAINTIFF ACACIA MEDIA TECHNOLOGIES CORPORATION'S MEMORANDUM OF POINTS AND AUTHORITIES IN OPPOSITION TO ROUND 3 DEFENDANTS' MOTION FOR SUMMARY JUDGMENT OF INVALIDITY UNDER 35 U.S.C. § 112 OF THE '992, '863, AND '702 PATENTS; AND SATELLITE DEFENDANTS' MOTION FOR SUMMARY JUDGMENT OF INVALIDITY OF THE '992, '863, AND '720 PATENTS

Hovsepian v. Apple, Inc.

Doc. 325 Att. 18

Dockets.Justia.com

A1.9

Estimation of Perceptual Entropy Using Noise Masking Criteria

James D. Johnston

AT&T Bell Laboratories

ABSTRACT

An estimate of the perceptual entropy of audio signals is created from the combination of several well known noise masking measures reported in the literature. Scharf's tone-masking-noise results, Hellman's results on noise-masking-tone and several individual's research on critical bands and spreading functions are combined with an heruistic method of tonality estimation to provide an estimate of the short term (64 mS) frequency masking templates for audio stimuli.

The perceptual entropy of each short term section of the audio stimuli is estimated as the number of bits required to encode the short term spectrum of the signal to the resolution required to inject noise at the masking template level. The numbers measured by this process provide an entropy estimate, for transparent coding, of 1.4 (mean) or 2.1 (peak) bits/sample for telephone speech (200-3200 Hz bandwidth sampled at 8 kHz). The entropy measures for audio signals of other bandwiths and sampling rates is also reported.

1. Introduction

Research on bit-rate reduction of speech signals is well established. Natural-sounding coding algorithms have reached a rate of approximately 16kB/s, and very mechanical sounding algorithms have reached 1kB/s. There seem to be two thresholds in current speech coding work. The first, at 16 kB/s upwards, depending on the talker, listener, and algorithm, is the threshold at which the listener can distinguish between the coded version and the original version. The second, at lower bit rates with coders more dependent on explicit source-modeling, is the rate at which the speech is perceived as "mechanical", "unnatural", or "reverberant". This threshold is currently at approximately 9.6kB/s, depending again on talker, listener and algorithm.

Although there are many coders operating below 16kB/s, none of these coders has yet provided a decoded signal that is transparent, i.e. one the listener cannot distinguish from the original. This apparent limit of 16kB/s or 2 bits/sample has persisted despite a considerable amount of work in the area of low bit rate speech coding. In spite of work that has produced speech of acceptable "communications quality" which offers little interference with the understanding of language at rates well below 16kB/s, the lack of transparency suggests that there may be a mechanism interfering with transparent coding of 200-3200 Hz speech below that rate.

The ear's perceptual mechanism is known to place a limit on the smallest spectral differences that can be discerned for tone masking noise [11] or noise masking tone [7], thus suggesting that a limit on the just audible difference (or coding error) in a more complex signal may be estimated, and a short-time bit rate estimated from this limit. This paper will discuss the estimated "Perceptual Bit Rate" or "Perceptual Entropy" (PE) that arises from that estimate.

The entropy estimates that were generated by this algorithm using the initial model based on [11], [7], and others are on the order of 1.5 bits/sample for telephone speech. This reinforces the limits that have empirically been suggested for speech coding, suggesting that the PE measurement may well estimate a limit for transparent bit rate reduction for signals presented to the human ear.

Signals at higher bandwidths (7, 15, or 20kHz) have not been investigated as much in terms of bit rate reduction. Current work [8] suggests that the neighborhood of 128kB/s for 15kHz bandwidth material more than suffices for transparent coding with fairly complex algorithms based on a frequency domain quantization that uses a masking model similar to that detailed in this paper. This bit rate corresponds to 4 bits/sample. Recent informal work suggests that 3 bits/sample for 15kHz signals is sufficient. Long-time averaged entropy estimates with this model are well below 4 bits/sample, ranging from .08 to 1.5 bits/sample depending on the input material, however the (short term) PE of one segment reaches 2.6 bits/sample.

2. Estimation of Perceptual Entropy

2.1 A Short History of Critical Bands and Masking Measurement

In the first studies of masking, [5] created the term "critical band". [4], in 1956, redefined "critical band" in terms of noise loudness. This redefinition was also shown to have significance in terms of the physical structure of the ear [6] [12]. [11], using the critical band method, provided a model for the masking thresholds for critical band noise masked by a single tone. [10], using results from Zwicker, provided a model to permit the modeling of masking for noise of an arbitrary bandwidth and a tone of arbitrary frequency. [7] provided a model for the masking thresholds for a single tone masked by critical band noise. The "spreading function" from [10] is equally applicable to Hellman's work.

In essence, critical band analysis has been shown to be very convenient for calculation of masking phenomena, hence analysis on the critical band, or "Bark" (after Barkhausen) scale is indicated for the calculation of the masking properties of signals that have spectral and temporal characteristics between those of tones and noise.

In short, the results in references [4-7][10-12] can be summarized as follows:

— Both tones-masking-noise and noise-masking-tones can be well modeled on the Bark scale.

— The spreading function provides a useful model for the interaction of masking signals that are not within a critical band.

— For tonal signals, the tones-masking-noise results, along with the spreading function, can calculate, in the Bark domain, a noise threshold at all frequencies.

— For noise-like signals, the noise-masking-tone results, again using the spreading function, can predict a tone-masking threshold at all frequencies.

While these are the results of special cases, i.e. pure tones or pure noise, they offer at least the extrema from which to interpolate the sensitivity of a listener to corruption of a complex signal containing both noiselike and tonal components.

2.2 The Need for Estimating Masking Thresholds for Complex Signals

When doing bit rate reduction of speech, music, or any arbitrary signal that will be presented to the auditory system, the object is to introduce either imperceptible or inoffensive degradation during the coding process. This process is naturally opposed to the idea of reducing the bit rate of the signal. Many subjective tests of speech coding algorithms have been reported [3] and at least one for "commentary grade" (6 kHz) speech and music [2]. In the report on the commentary grade signals, the results were quite complex, with each of the better coders being preferred overall for some subset of the input signals. Unfortunately, each of the better coders was also highly unpreferred for some other subset of the signals. Several sensitivities are described in that paper. [9] shows a statistical analysis of the signals with atypical behavior, and identifies some of the problems that particular coders had with particular signals.

One of the conclusions of [9] was that the masking properties of the various input signals had to be carefully considered. In particular, one example was that of several signals that had been high-pass filtered. A long-term critical band analysis showed that the high-pass filtered signals were particularly sensitive to relatively small amounts of low frequency quantizing noise. Although the SNR of those particular coded signals was well above that which would usually be regarded as good, the particular spectrum of the input (masking) signal resulted in a noise loudness that was considered offensive by most listeners.

2.3 The measurement of Perceptual Entropy

The masking threshold for a signal indirectly shows the amount of quantization that may be applied in the frequency domain, i.e., the quantization, according to the masking model, that may be done without corrupting the signal such that it can be distinguished from the original. The part of the signal that can be changed without making the signal distinguishable is therefore perceptually redundant, and the part that must be reproduced represents real information that can be quantized and measured.

In an ideal transform coder, the step size and number of levels in the quantizer for each line could be set independently, and with no side information to communicate the level and/or bit allocations to the decoder. If the step size in this ideal coder were set such that the total noise injected at each frequency corresponds to the threshold, i.e. the minimum number of quantizer levels are used, then the number of bits required to encode the entire transform represents an estimate of the minimum number of bits necessary to transmit that block of the signal. The total rate, divided by the number of samples coded, represents the

CH2561-9/88/0000-2524 $1.00 © 1988 IEEE

Exhibit R Page 373

per-sample rate. The per-sample bit rate of this ideal transform coder is called the "Perceptual Entropy" of the signal. A block diagram of this process is shown in Fig. 1.

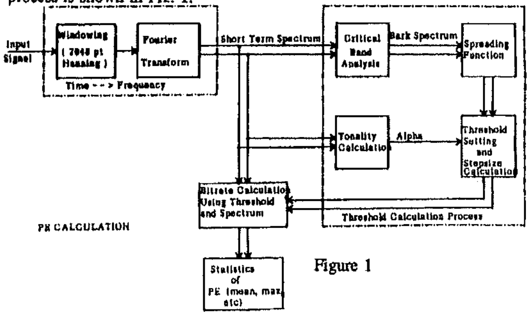

Input Signal
Windowing (2048 pt Hanning)
Fourier Transform
Time --> Frequency
Short Term Spectrum
Critical Band Analysis
Bark Spectrum
Spreading Function
Tonality Calculation
Alpha
Threshold Setting and Stepsize Calculation
Bitrate Calculation Using Threshold and Spectrum
Threshold Calculation Process
PE CALCULATION
Statistics of PE (mean, max, etc)

Figure 1

This model is particularly attractive, as it takes into account all of the artifacts and redundancies in the input signal to the same extent that the ear does, in essence taking care of pitch (for speech), short term spectral model, etc. without explicit calculation.

2.4 Details of the PE Algorithm

As shown in Fig. 1 there are three main parts to the algorithm. They are:

— Windowing and transformation to the frequency domain.

— Calculation of the masking threshold.

— Calculation of the number of bits required to quantize the spectrum of the signal.

2.4.1 Windowing and FFT The windowing and frequency transformation are done by a standard Hanning window followed by a real-complex FFT of length 2048. The first 1024 complex lines are retained (including the DC and π lines counted as one line).

2.4.2 Calculation of the Masking Threshold There are several steps involved in calculating the masking threshold. They are:

— Critical Band Analysis of the Signal

— Applying the Spreading Function to the Critical Band Spectrum

— Calculating the Spread Masking Threshold

— Accounting for Absolute Thresholds

— Relating the Spread Masking Threshold to the Critical Band Masking Threshold

2.4.2.1 Critical Band Analysis We are presented with the spectrum $Re(\omega)$,$Im(\omega)$ of the signal from the FFT. The complex spectrum is converted to the power spectrum, $P(\omega)=Re^2(\omega)+Im^2(\omega)$. The spectrum is then partitioned into critical bands according to [11], and the energy in each critical band summed, i.e. $B_i=\sum_{\omega=bl_i}^{bh_i} P(\omega)$ where bl_i is the lower boundary of critical band i, bh_i is the upper boundary of critical band i, and B_i is the energy in critical band i, where i=1 to i_{max}, where i_{max} is dependent on the sampling rate. Fig. 2a shows a power spectrum and critical band spectrum for 64ms of a loud brass passage. The bark spectrum in Fig. 2a is the staircase plotted above the frequency spectrum. The horizontal scale of the figure is (Hz) frequency, rather than barks, so the critical band spectrum shows the widening of critical bands at high frequencies as well as the energy in the critical band.

A true critical band analysis would sum $P(\omega)$ at each ω in order to create a continuous critical band spectrum. For the purposes of the PE calculation, the discrete critical band represents a close approximation.

B_i is then passed to the spreading function.

2.4.2.2 Spreading Function The spreading function, as given in [10] is used to estimate the effects of masking across critical bands. The spreading function is calculated for $abs(j-i)\le 25$, where i is the bark frequency of the masked signal, and j the bark frequency of the masking signal, and placed into a matrix S_{ij}. The convolution of the $B(\omega)$ with the spreading function is implemented as a matrix multiplication, i.e. $C_i=S_{ij}*B_i$. The value of C_i denotes the spread critical band spectrum. Fig. 2b shows the results of spreading of the bark spectrum in Fig. 2a. Looking at Fig's 2a and 2b, the most obvious effect of the spreading function is the spreading of the peak in critical band 7 toward critical band 8, which greatly raises the threshold energy in band 8. Another visible effect is the elevation of the threshold at high frequencies due to masking by lower frequency parts of the spectrum.

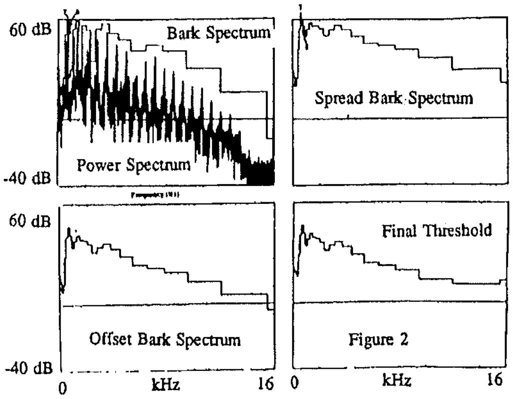


Figure 2

2.4.2.3 Calculating the Noise Masking Threshold There are two noise masking thresholds discussed in [11] and [7]. The first, for tone masking noise, is estimated as $14.5+i$ dB below C_i, where i is the bark frequency. The second, for noise masking a tone, is estimated as 5.5 dB below C_i uniformly across the critical band spectrum.

In order to determine the noise-like or tone-like nature of the signal, the Spectral Flatness Measure (SFM) is used. The SFM is defined as the ratio of the geometric mean (Gm) of the power spectrum to the arithmetic mean (Am)of the power spectrum. In this use, the SFM is converted to dB, i.e. $SFM_{dB}=10\log_{10}\frac{Gm}{Am}$, and further used to generate a coefficient of tonality, α, as follows:

$\alpha=\min(\frac{SFM_{dB}}{SFM_{dB\max}},1.)$ An SFM of SFM_{max}=-60 dB is used to estimate that the signal is entirely tonelike, and a SFM of zero dB to indicate a signal that is completely noiselike. In other words, a SFM of -30dB would result in α=.5 , and a SFM of -75dB results in α=1.000 .

The offset (O_i) in dB for the masking energy in each band i is then set as

$O_i=\alpha(14.5+i)+(1-\alpha)5.5$. In other words, the index α is used to geometrically weight the two threshold offsets.

The threshold offset is then subtracted from the spread critical band spectrum to yield the spread threshold estimate, T_i : $T_i=10^{\log_{10}(C_i)-\frac{O_i}{10}}$.

In practice the use of the SFM to estimate tonelike signals is fairly accurate, as most tonelike signals such as organ, sine waves, flute, etc have an SFM that is close to or over the limit of $SFM_{dB\max}$, and signals such as percussion have SFM's that are between -5 and -15dB. Speech signals at 200-3200 Hz are in the range [9] of -20 to -30 dB. Fig. 2c shows the plot of the spread threshold estimate for the data in Fig. 2a and 2b. The obvious difference in Fig. 2c is the much lower absolute level of the threshold. The tilt due to O_i is harder to see, even though it is present.

2.4.2.4 Renormalizing the Spread Spectrum Due the the fact that the spectrum spreading functions do not have a normalized gain, and that the gains in the critical bands around zero and the sampling rate will be different, the spread spectrum is renormalized by $\frac{1}{the\ DC\ Gain}$ for each critical band.

2.4.2.5 Including Absolute Threshold Information After the noise energy is renormalized in the bark domain, the bark thresholds are compared to the absolute threshold measurements due to [5]. Since the masking thresholds have thus far been calculated without reference to absolute level, they must be checked to make sure that they do not demand a level of noise below the absolute limits of hearing.

The absolute thresholds are set such that a signal at 4kHz, with a peak magnitude of ±1 least significant bit in a 16 bit integer is at the absolute threshold of hearing. Any critical band that has a calculated noise threshold lower than the absolute threshold is changed to the absolute threshold for that critical band. At high and low frequencies, the absolute threshold varies inside the critical band. In such cases, the mean of the critical band edges is used.

Fig. 2d plots the final threshold, after renormalization and adjusting for absolute threshold conditions. This is the threshold used to actually calculate the bit rate. The effects of the threshold are visible at both ends of the spectrum, where the absolute thresholds are above the calculate threshold due to the very low energy of the signal at those frequencies.

Exhibit R Page 374

2.4.3 Calculation of the PE The PE is calculated by measuring the actual number of quantizer levels to follow the signal in the frequency domain, given a step size in the quantizer that will result in noise energy equal to the audibility threshold. Remember that T'_i is in the power domain, and that the quantization energy must be spread across k spectral lines in each critical band. It is assumed in the absence of a much more detailed critical band model that the noise is spread equally across the entire band. The assumption that the distribution of quantization error is uniform in the amplitude domain leads to a noise energy equal to $\frac{\delta^2}{12}$.

The step size T''_i is calculated as follows:

First, the energy must be spread across the entire critical band, i.e. the energy at each spectral frequency $=\frac{T'_i}{k_i}$. Then, since the real and imaginary parts of the spectrum are quantized independently, the energy at each frequency must be divided by 2, i.e. the energy at each spectral component $=\frac{T'_i}{2k_i}$. The energy due to quantization is $\frac{\delta^2}{12}$ so $\frac{\delta^2}{12}=\frac{T'_i}{2k_i}$ or, since $\delta=T''_i$; $T''_i=(\frac{6.T'_i}{k_i})^{\frac{1}{2}}$ where T'' is the quantizer step size.

This is done in each of the j critical bands: $N_{Re}(\omega)=abs(nint(\frac{Re(\omega)}{T''_i}))$ and $N_{Im}(\omega)=abs(nint(\frac{Im(\omega)}{T''_i}))$ for each ω within critical band i. *Abs* represents the scalar absolute value function, and *nint* a function that returns the nearest integer to its argument. N_* represents the actual (integer) quantized value of each line.

Then, for each ω, and individually for real and imaginary parts, $N_{(Re\ or\ Im)}(\omega)$ is tested as follows:

— If $N_{(Re\ or\ Im)}(\omega)=0$, $N'_{(Re\ or\ Im)}(\omega)=0$.

— If $N_{(Re\ or\ Im)}(\omega)\neq 0$, then $N'_{(Re\ or\ Im)}(\omega)=\log_2(2.N_*(\omega)+1)$

This operation assigns a bit rate of zero bits to any signal with an amplitude that does not need to be quantized, and assigns a bit rate of $\log_2$(*number of levels*) to those that must be quantized. For example, if the integer number is one, three levels (-1 , 0 , +1) are required to quantize the particular line. As the signs of the various spectral lines are random, the sign information must be included. When no levels are necessary, transmission of the sign bit is unnecessary, therefore a 0 is assigned to that line.

The total bit rate is then calculated as $Total\ Rate=\sum_{\omega=0}^{\pi}(N'_{Re}(\omega)+N'_{Im}(\omega))$, and the rate per sample calculated as $PE=\frac{Total\ Rate}{2048.}$.

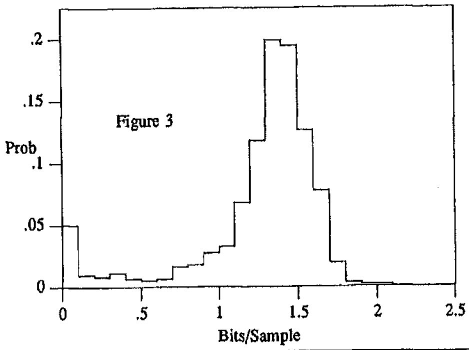
The term PE is used throughout the following discussion to indicate the 2048 sample perceptual entropy, regardless of sampling rate, bandwidth, etc. While this window length is not the best suited for low (8kHz, 14kHz) sampling rates, interpolation of those signals to 32 or 28 kHz has shown that the PE measurements do not vary enormously. The variability (block to block) increases as the effective window length decreases, but the mean and extrema do not change significantly.

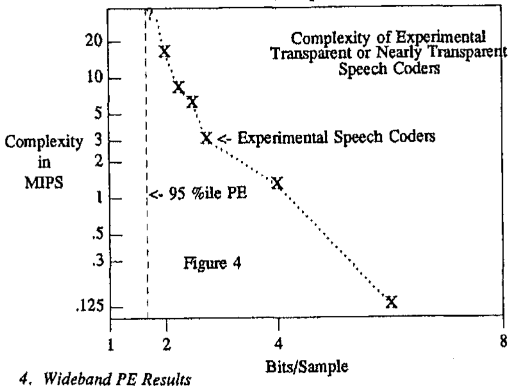
3. PE Measurements for 8kHz speech

Fig. 3 shows a histogram of the PE measurement for a large selection of speakers, constituting about 1 hour of telephone bandwidth speech. The statistics of the PE measurement show that there is a small but significant difference in bit rate between the mean and the peak rate across 2048 sample blocks, suggesting that use of the non-stationarity in time would lead to about .3 bit/sample (or 2.4kB/s) savings. It appears that for the speech samples that have been analyzed, delays of up to a 1 to 2 seconds are necessary to make effective use of the time non-stationarity. The three utterances, as analyzed, include short silences at the beginning and end that are included in the mean and minima measurements. The PE measurement for telephone bandwidth speech currently includes about 1 hour of male, female, and child utterances.

Currently, perception of the quality of transparent speech is based on the worst case condition, the peak rate, as time domain bit allocation over a time window of greater than 64ms has yet to be constructively used. [2] The tested material has a peak rates of 2.1 bits/sample, and a 95th %'ile of 1.65 bits/sample.

Fig. 4 shows a graph for the complexity of various coders that are known to provide high quality coding of telephone speech. The bit rates in Fig. 4 are estimates for some coders, as some of the low rate coders have not been traditionally designed to provide complete transparency at a bandwidth of 200-3200Hz.





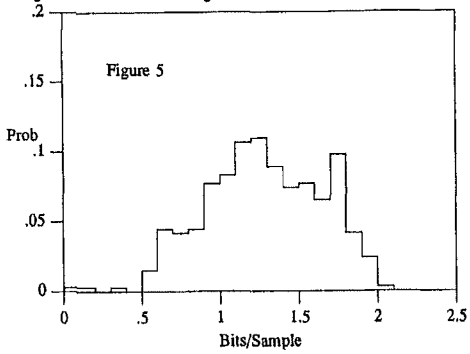
Complexity of Experimental
Transparent or Nearly Transparent
Speech Coders
Complexity
in
MIPS
20
10
5
3
2
1
.5
.3
.125
X <- Experimental Speech Coders
<- 95 %ile PE
Figure 4
1
2
4
8
Bits/Sample

4. Wideband PE Results

The PE measure was run on a set of signals with 7 and 15 kHz bandwidth, sampled at 14kHz and 32kHz respectively. There were 8 signals available for the 7kHz material, and 15 for the 15kHz material. As detailed below, all of the PE estimates were in the range of 0.1 to 2.1 bits/sample.

4.1 Signals with 7 kHz Bandwidth

Fig. 5 is a histogram of the PE measurement for a set of the 7kHz signals. As before, the peak estimates place a bound on the minimum coding rate with current coding methods.



Several of the signals have significant silence at the beginning and end. The very low PE estimates occur when the signal is nearly below the absolute thresholds as estimated in Section 2.4.2.4. The other signals are selected to be the middle of active passages, and as a result the minima are roughly .5 bit/sample.

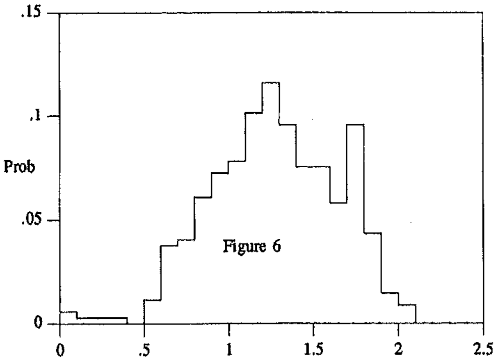
The maxima of several signals reaches 2.1 bits/sample, suggesting that 2.1 bits/sample is a good estimate of a minimum rate. Currently, coders of medium complexity can provide transparent coding at greater than 4 bits/sample for this bandwidth.

Exhibit P Page 375

4.2 Signals with 15kHz Bandwidth

Fig. 6 summarizes the results of PE measurements made on a wide variety of 15kHz music signals. The range of the maxima for this set of data is .35 to 2.1 bits/sample. The five signals with the highest and very similar maxima are quite varied in source and style, with the highest being a percussion performance, and the others being an electronic pop/jazz performance, an a-capella vocal, a baroque harpsichord, and a "heavy metal" rock excerpt.

Again, the means of the various signals suggest that a coding algorithm taking advantage of very long time non-stationarity can provide at least



0.5 bit/sample reduction in rates, however the length of time required for this time buffering is estimated at 2-4 seconds as the signals tend to vary fairly slowly in entropy, with a highly entropic section lasting up several seconds, followed by a lower entropy section, and so on.

The currently available coding methods for 15 (or 20) kHz audio have reached between 3 and 4 bits/sample [8] with transform coding techniques, suggesting that further gains in coding efficiency may be either tied to long time delays or efficient extraction of stereo information. In any case, the PE measurements suggest that a limit of around 2 bits/sample is likely for monophonic material.

5. Conclusions

The PE measurement is based on the hearing mechanism, and the technique is essentially independent of the signal source. The absence of source analysis, the accuracy, and the direct method of calculation make the masking model used for the PE measurement a useful approximation for both estimation of the perceptual entropy and for use in coding. The method of estimation for complex signals, using the SFM as a measure of tonality and the geometric interpolation of the thresholds based on the tonality estimate, provides an estimate of the noise threshold that suffices for use in the less than 4 bit/sample coder for 15kHz signals reported in [8].

The PE measurements for 200-3200Hz speech provide a bound that is compatible with the performance of currently available coding methods, and a plot (Fig. 4) of the complexity of coders that approach that bound seems to indicate some sort of assymtotic increase in complexity. As speech has been the most carefully investigated signal in terms of bit rate reduction, the corraboration of the PE numbers with speech experience is very positive.

5.1 Caveats

This method of estimating PE does not take into account all source redundancy, nor does it expressly include "forward" masking. Most important in the case of highly stationary signals is that of redundancy that extends to considerably more than 64 ms. If a particular signal, for example a sine wave, is very slowly varying such that its spectrum can be accurately extraplolated, the PE estimate will be high. Very few physical systems or stimuli that are presented to the ear are in this catagory.

Another noticible way that PE measurement can result in an overestimate is in the case of text-to-speech, as the PE measurement does not take into account the contraints that are built into the text-to-speech algorithm. In such cases, the PE does represent the bit rate that can be achieved without knowledge of the generating system. In general, where a process contains no physical uncertainty and has specific and very precise rules, the PE will provide an overestimate because the process is constrained in ways unknown to the PE measurement.

In general, for signals generated by acoustic processes, the PE will be a good estimate to the lower bound of transparent encoding.

5.2 Implications for Future Coding Work

If the conclusions reached by this method of PE calculation are correct, the current state of the art in speech coding is rapidly approaching its lower bound in bit rate. The way to improve the performance of coders below these bounds will be to develop methods of impairing signals that are not objectionable, even though they are audible, rather than trying to make coding algorithms transparent, or to generate much more efficient representations of the speech generation mechanism. The process of making non-objectionable distortions has already begun as many current standard speech algorithms are considered to produce acceptable speech that is clearly distinguishable from the original, when level, bandwidth, etc are carefully controlled. Some of these "communications quality" algorithms reach bit rates well below the calculated limit of transparency.

The implications for wideband work indicate that coding algorithms have yet to reach the estimated limits of coding efficiency. Several algorithms, among them [8], have reduced the bit rate to near or at 3 bits/sample, but the absolute limit of all the signals as estimated by the PE is near 2.

6. Acknowledgements

The author thanks J. L. Hall for both his continued help with the perceptual models and his patience in providing references, tables, missing information and more, and the researchers cited in the references for their essential results.

References

1. - R.V. Cox, J.H. Snyder, R.E. Crochiere, D.E. Bock, J.D. Johnston, "Testing of Wideband Digital Coders", *Proceedings of ICASSP 1984*, pp 19.3.1-19.3.4.
2. - "Cox, R.V., and Crochiere, R.E., "Multiple User Variable Rate Coding for TASI and Packet Transmission Systems", *IEEE Transactions on Communications*, Vol. Com-23, #3, March 1980, pp334-344.
3. - W.R. Daumer, "Subjective Evaluation of Several Effecient Speech Coders", *IEEE Transactions on Communications*, April 1982, pp655-662.
4. - Feldkeller, R., and Zwicker, E. *Das Ohr als Nachrichtenempfanger*, Stuttgart: Hirzel, 1956.
5. - Fletcher, H, Auditory Patterns, *Reviews of Modern Physics*, 1940, 12, pp47-65.
6. - Greenwood, D.D., Critical bandwidth and frequency coordinates of the basilar membrane, *Journal of the Acoustical Society of America*, 1961, 33, 1344-1356.
7. - Hellman, R. P., "Assymmetry of Masking Between Noise and Tone, *Perception and Psychophysics*, pp 241-246, 1972.
8. - J. D. Johnston, "Transform Coding of Audio Signals using Perceptual Noise Criteria", IEEE Journal on Selected Areas in Communications, to be published Feb. 88.
9. - J. D. Johnston, Abstract to "Digital Coding of Musical Sound - Some Statistics of Interest", IEEE-ASSP 1986 Mohonk Conference on Digital Audio.
10. - Schroeder, M.R., Atal, B. S., Hall, J. L., Optimizing Digital Speech Coders by Exploiting Masking Properties of the Human Ear,
11. - From *Foundations of Modern Auditory Theory*, edited by Jerry V. Tobias, Academic Press, NY, NY, Chapter 5, by B. Scharf.
12. - Zwislocki, J., Analysis of some auditory characteristics, from *Handbook of Mathmatical Psychology*, edited by R.D. Luce, R. R. Bush and E. Galanter, NY, Wiley 1965. J. Acoust. Soc. Am. 66(6), Dec. 1979, pp1647-1651.

Exhibit R Page 316